A MODIFIED GAUSS-NEWTON ALGORITHM AND NINETY-SIX WELL MICRO-TECHNIQUE FOR CALCULATING MPN USING EXCEL SPREADSHEETS[1]

P. IRWIN[2], S. TU, W. DAMERT and J. PHILLIPS

USDA, Agricultural Research Service
Eastern Regional Research Center
Microbial Biophysics & Biochemistry Laboratory
600 East Mermaid Lane
Wyndmoor, PA 19038

Accepted for Publication July 24, 2000

ABSTRACT

Conventional most probable number (MPN) methods seek a calculated value for microbial concentration (Δ, mL^{-1}) which induces the total binomial probability function (P_{total}) to approach its maximum limit. In fact, such techniques are the only statistically compelling procedures available for determining MPN when utilizing a small set of observations per dilution (e.g., n = 3-8). However, as n approaches a large value, statistical routines which assume a normal distribution might be applied to ascertain the MPN. With this in mind, we produce herein a modified Gauss-Newton "linearization" (curve fitting) algorithm for determining Δ (n = 96) from binomial micro-plate assays which are readily automated using 96-well micro-plate readers. This technique, an iterative protocol, is less cumbersome than many traditional MPN procedures and has certain advantages. Data derived from this method were not only close to MPN estimations using a direct technique based on the conventional maximum probability resolution (MPR) concept but also displayed more favorable chi-squared (χ^2) statistics.

[1]Reference to a brand or firm name does not constitute endorsement by the U.S. Department of Agriculture over others of a similar nature not mentioned.

[2]To whom correspondence should be addressed; Tel: 215-233-6420; E-mail: pirwin@arserrc.gov

INTRODUCTION

For many microbiological applications it is only feasible to estimate microbial concentration using the dilution method (Barkworth and Irwin 1938; Dickson 1989; Best 1990; Blais and Yamazaki 1991; Turpin *et al.* 1993). Specifically, the dilution method consists of taking a sample from a liquid source, making serial dilutions from it, introducing an aliquot of each of the dilutions into an appropriate culture medium, incubating samples at a suitable temperature, and observing if any growth (e.g., number of positive samples, p, out of n observations) occurs. One of the most critical steps in the satisfactory application of the dilution method is obtaining a statistically cogent estimation of the "most probable number" of organisms (MPN) in the original sample. The MPN concept is quite venerable inasmuch as its formal description was first made in 1915 (McCrady 1915). McCrady's reckoning of MPN is based on the application of binomial probability theory, and it depends upon 2 primary assumptions: microbes are distributed randomly throughout the test fluid and growth occurs if the aliquot used contains one, or more, organisms (McCrady 1915; Cochran 1950). In 1969, Fung and Kraft miniaturized the dilution method using "microtiter" plates in an 8 dilution, 3 replicates (n = 3) per dilution MPN format whereas Rowe *et al.* (1977) expanded this procedure to 12 dilutions and 8 replicates (n = 8) per dilution. Computer-based computational methods (Best 1990; Briones and Reichardt 1999) for the estimation of MPN are similar, in principle, to earlier efforts (McCrady 1915; Halvorson and Ziegler 1933) which are the *only* statistically valid procedures available for determining MPN when utilizing a small set of observations (n) per dilution (Rowe *et al.* 1977; Haines *et al.* 1996; Humbert *et al.* 1997).

In this manuscript, we developed a nonlinear least squares method for determining MPN when n = 96 using micro-plate turbidity assays, and quantitatively compared this technique with a conventional MPN estimation from the same data. We have arbitrarily chosen n = 96 (8 × 12 wells per plate) since this is the standard plate format with micro-plate readers. Thus, each 96-well plate represents one dilution. Also, 96 wells × 50 μL per well gives 4.8 mL of sample used per dilution and is similar to the traditional 5 tube multiple dilution MPN (de Man 1975, 1983). The more observations per dilution one makes the greater the accuracy of the calculation. Most curve fitting procedures (Hartley 1961; Draper and Smith 1980) are somewhat tedious to perform as matrix transposition, multiplication, and inversion are involved. However, because our method involves only a one parameter fit (e.g., solving for the MPN), the algorithm is more straightforward than many traditional MPN computational procedures (Best 1990; Briones and Reichardt 1999) and has some compelling advantages.

MATERIALS AND METHODS

General

All bacteria used in this research were *Salmonella enteritidis* (avian isolate from Dr. K. Rajkowski [USDA, ARS, Eastern Regional Research Center]; identification confirmed by comparisons of ribosomal DNA with known *S. enteritidis* isolates and various biochemical tests). Brain-heart infusion agar (BHIA) and broth (BHI) were obtained from Difco (Detroit, MI). Rainbow agar (RA), selective for hydrogen sulfide-producing strains of *Salmonella*, was obtained from Biolog, Inc. (Hayward, CA). After plating onto BHIA and RA (6 plates each) with an Autoplate 7000 "spiral" plating apparatus (Spiral Biotech; 50 µL per 10 cm plate) and incubating for 16-18 h at 37C, total aerobic plate count cell density (δ) was determined by manual colony counting. We utilize the terms δ and Δ which represent measures of solution cell density. However, δ is based on petri plate colony counting and has dimensions of either mL^{-1} or colony forming units mL^{-1} ($CFU\ mL^{-1}$). The equivalent term (Δ) derived from binomial enumeration methods has units of either $MPN\ mL^{-1}$ or mL^{-1}. Both δ and Δ should numerically agree ($\pm$ 10-15%) for pure cultures. Typically only serial dilutions estimated to contain 1,000-2,000 $CFU\ mL^{-1}$ were utilized for total aerobic plate count (e.g., 50-100 CFU per plate). Phosphate-buffered saline (PBS; 100 mM, 0.85% [w/v] NaCl, pH 6.7) was used as the diluent throughout.

Ninety-Six Well Micro-Plate Assays

"Microtiter" or micro-plate turbidity assays (Irwin *et al.* 2000; n = 96) were used as our binomial method. For each *Salmonella* dilution to be enumerated, a 150 µL aliquot of sterile BHI was pipetted into every well of disinfected 96-well micro-plates (one plate per dilution) and 50 µL of inoculum introduced (PBS alone or PBS-diluted cells; 0-70 $CFU\ mL^{-1}$). All plating was performed in a Type II Vertical Microbiological Hood. Inoculated plates were stored in a sterile container and incubated overnight at 37C. After 16-18 h, the number of positive responses (p), based on turbidity, was recorded. Visual and optical density (OD on a Perkin-Elmer HTS7000+ plate reader) turbidity determinations produced identical results. Figure 1 displays optical density (λ = 492 nm) as a function of incubation time for 3 representative positive sample wells inoculated with 50 µL of ca. 3 $CFU\ mL^{-1}$ *S. enteritidis* whereupon the time to approximately half-maximal optical density (β) was between 8-9 h postinoculation ($\Delta OD_{492\ nm} \sim 0.4$). The inset of Fig. 1 displays the optical density dependence upon bacterial concentration (i.e., standard curve) where the limit of detection was approximately 10^8 $CFU\ mL^{-1}$. Thus, presence or absence of turbidity is determinable with a micro-plate reader in ca. 9 h postinoculation. Using this technique, and the computations which follow, we observe a limit of detection of about 0.4 mL^{-1} (based on Appendix).

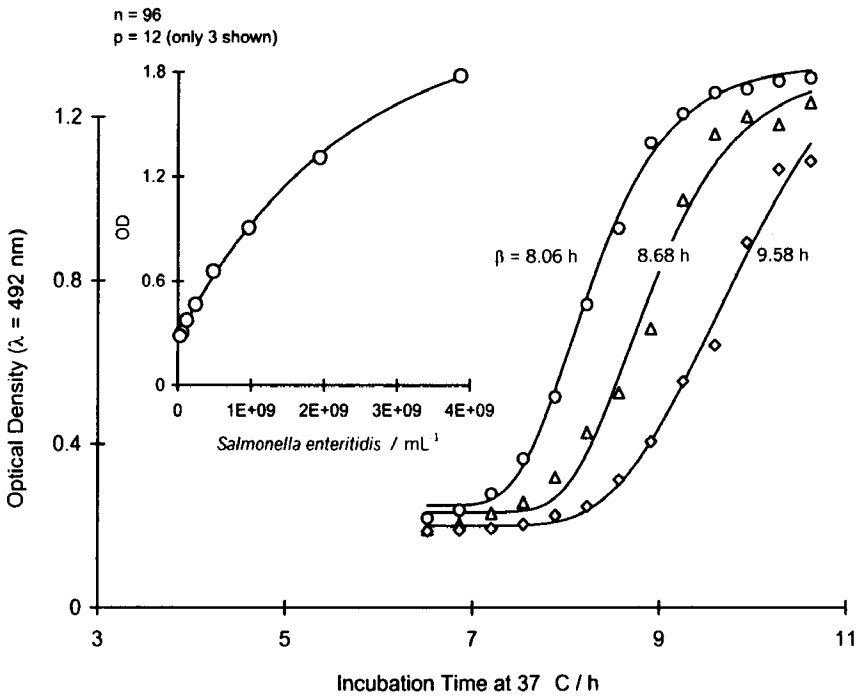


FIG. 1. TIME DEPENDENCE OF *SALMONELLA ENTERITIDIS* TURBIDITY AS DETERMINED BY OPTIAL DENSITY (OD, λ = 492 NM) USING A 96-WELL MICRO-PLATE READER (T = 37C)

Upon sufficient dilution, assuming a random sampling, our 96-well micro-plate assay allows only two possible outcomes: the volume (v) to be tested either does (p = positive) or does not contain the organism and represents, by definition, a binomial population (Steel and Torrie 1960). Thus, the probability, P_i, (McCrady 1915; Halvorson and Ziegler 1933) of observing p_i positive responses out of n samples after inoculating each with some volume (v_i) is

$$P_i = \frac{n!(e^{-v_i \Delta})^{n-p_i}(1-e^{-v_i \Delta})^{p_i}}{(n - p_i)!\,p_i!} \qquad (1)$$

By normalizing the derivative of P_i with respect to Δ and setting this equal to 0 it is possible to find the value of Δ that corresponds to the maximum value of P_i (e.g., the MPN; Best 1990; the symbolism, $\partial_x y = \frac{\partial y}{\partial x}$, is utilized throughout this work)

$$\frac{\partial_{\Delta} P_i}{P_i} = \left(p_i - n + \frac{p_i}{e^{v_i \Delta} - 1} \right) v_i = 0 . \tag{2}$$

Solving Eq. 2 for p_i yields

$$p_i = n - e^{-v_i \Delta} n . \tag{3}$$

Since a sequence of dilutions, each of volume v_i, is equivalent to the product of a fixed volume, v, and dilution factor Φ_i

$$v_i = v\Phi_i \tag{4}$$

then

$$p_i = n - e^{-v \Phi_i \Delta} n . \tag{5}$$

The experimental variable Φ_i,

$$\Phi_i = \prod_{h=1}^{h=(i-1)} \frac{V_{(i-h)}}{V_{(i-h)} + V_{diluent}}, \tag{6}$$

is a coefficient which defines each i^{th} dilution; for example, the dilution factor for the 11^{th} sample in a serial dilution ($\Phi_0 = 1$ for the initial, or undiluted, sample) is

$$\Phi_{11} = \frac{V_{10}}{V_{10} + V_{diluent}} \times \frac{V_9}{V_9 + V_{diluent}} \times \cdots \times \frac{V_1}{V_1 + V_{diluent}}; \tag{7}$$

v_i (Eq. 7) is the volume taken from the i^{th} sample utilized to make the $\{i + 1\}^{th}$ dilution by adding $v_{diluent}$ of buffer. Thus, Eq. 5 quantitatively describes the relationship of p_i as a function of Φ_i at fixed v and Δ.

Maximum Probability Resolution Methods (MPR): Traditional MPN Calculation on Microsoft Excel®

MPN calculations can be performed using one of two schemes. The direct method involves calculating the full array of $P_{total}(= \Pi_i P_i)$ as a function of Δ and finding the value of this variable which corresponds to the maximum in P_{total} (i.e., the MPN). Alternatively (Best 1990, Briones and Reichardt 1999), one can take advantage of the fact that Δ approaches the MPN where $\partial_\Delta P_{total}$, or some related function (Eq. 2), approaches zero. For example, the single dilution $\partial_\Delta P$ or $\partial_\Delta P/P$ (Eq. 2) equations, when set to zero, can be easily solved for Δ (Halvorson and Ziegler 1933)

$$\Delta_i = \frac{\mathrm{Log}_e \frac{n}{n - p_i}}{v \Phi_i}. \tag{8}$$

Of course, multiple dilution MPN calculations are more involved than the single dilution case since Δ, the cell density of the starting dilution ($\Phi_1 = 1$), can not be evaluated directly. To overcome this limitation, using the direct MPR method as an example, final values for MPN (Table 1) were determined with an Excel protocol which performs a calculation that picks the MPN directly from a large total probability array: $P_{total} = P_1 \times P_2 \times \ldots \times P_I = \Pi_i P_i$ calculated as a function of Δ (1000 Δ points per P_i); Fig. 2 defines the approximate 'start' and 'end' limits. Thus, a large (for 3 dilutions, 440 Kb) spreadsheet was created and an Excel function utilized to pick the maximum value in P_{total} ($[P_{total}]_{max}$)

$$= \mathrm{MAX}(\mathbf{P}_{total}) \tag{9a}$$

and then look up

$$=\mathrm{MATCH}([P_{total}]_{max}, \mathbf{P}_{total}, 0) \tag{9b}$$

the row containing the Δ value associated with $[P_{total}]_{max}$ which is the MPN (bold characters represent arrays of calculations or numbers within an Excel worksheet). Table 1 compares MPN values from this MPR method with those obtained from a well-known three dilution MPN Table (de Man 1983). These data show close agreement (average absolute value deviation was ca. 2% mostly due to rounding-off differences) between the MPR protocol and de Man's calculations. The MPR Δ_{final} (or MPN) error term was calculated as

$$\sigma_{MPR} = \frac{1}{\sqrt{\sum_{i=1}^{I} \frac{p_i \left(v\Phi_i\right)^2 e^{v\Phi_i \Delta_{final}}}{\left(e^{v\Phi_i \Delta_{final}} - 1\right)^2}}}. \qquad (9c)$$

TABLE 1.
3 DILUTION MPNS: TABULAR (DE MAN 1983) AND THE MAXIMUM PROBABILITY RESOLUTION (MPR) METHOD

			MPN (n = 5)		
p_1[a]	p_2	p_3	*Table*	*MPR*	± σ
5	5	0	24	24.00	13.5
5	4	4	35	34.50	14.61
5	4	3	28	27.81	12.24
5	4	2	22	22.08	10.19
5	4	1	17	17.22	8.46
5	4	0	13	13.00	6.97
5	3	2	14	14.04	6.58
5	3	1	11	10.88	5.63
5	3	0	8	7.91	4.60
5	2	2	9	9.45	4.79
5	2	1	7	7.00	3.96
5	2	0	5	4.93	2.92
5	1	2	6	6.30	3.45
5	1	1	5	4.56	2.59
5	1	0	3	3.30	1.86
5	0	1	3.1	3.14	1.72
5	0	0	2.3	2.31	1.28
4	4	0	3.4	3.35	1.41
4	3	1	3.3	3.26	1.36
4	3	0	2.7	2.71	1.19
4	2	1	2.6	2.65	1.15
4	2	0	2.2	2.16	0.99
4	1	1	2.1	2.11	0.97
4	1	0	1.7	1.69	0.83
4	0	1	1.7	1.66	0.81
4	0	0	1.3	1.27	0.68
3	3	0	1.7	1.72	0.75

[a] $v = 1$, $\Phi_1 = 1$, $\Phi_2 = 0.1$, $\Phi_3 = 0.01$

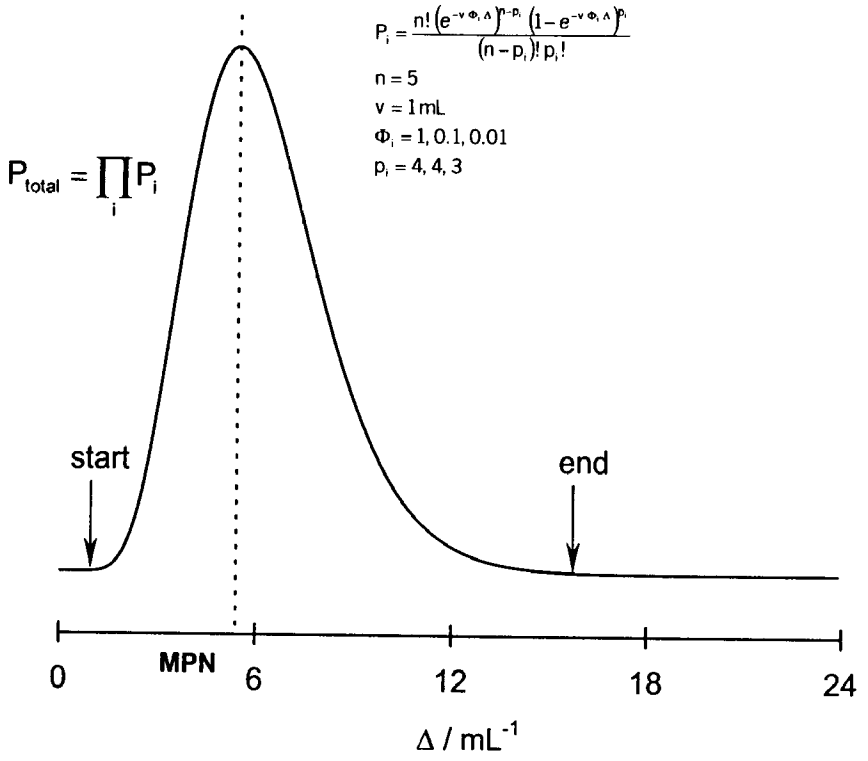


FIG. 2. DEFINITION OF THE BEGINNING AND ENDING LIMITS FOR THE TOTAL BINOMIAL PROBABILITY DISTRIBUTION FUNCTION ($P_{total} = \Pi_i P_i$) UTILIZED FOR THE DIRECT DETERMINATION OF MPN

RESULTS AND DISCUSSION

All MPN methods with which we are familiar with either utilize some mathematical manipulation of Eq. 2 in order to locate the value of Δ which induces the *total* probability distribution function (P_{total}) to approach its maximum limit, or directly calculate P_{total} and search for the maximum probability and associated MPN. Such MPR methods are the *only* statistically persuasive protocols available for determining MPN for small n (Rowe *et al.* 1977; Haines *et al.* 1996; Humbert *et al.* 1997). The benefit of numerically determining MPN in this way, over using a table (de Man 1975, 1983), is that any number of dilutions or assay volumes can be utilized within each interdependent data set thereby gaining greater experimental flexibility (Briones and Reichardt 1999). However, as n approaches a large number (e.g., n = 96), the P distribution with Δ becomes almost Gaussian or normally-

distributed. For instance, Fig. 3 displays plots of $\partial_\Delta P$ and P (single dilution; A: n = 8; B: n = 96; p = n/2 in both; all data normalized) as well as the best-fit Gaussian based on nonlinear regression of 26 evenly-spaced points from each P array. In Fig. 3A, P (n = 8) is quite asymmetrical as evidenced by its skewed first derivative ($\partial_\Delta P$) and the Gaussian does not fit P well. However, when n = 96 (Fig. 3B) $\partial_\Delta P$ is nearly symmetrical and the Gaussian fit is similar to P. The nearly-symmetrical binomial probability distribution at large n, as well as a necessary abatement of rounding-off and sampling errors, might imply that methods which assume a normal distribution could be used to determine MPN. If this is true, we should be able to directly fit binomial data (p_i), collected as a function of various dilutions (Φ_i), to an appropriate function (Eq. 5) using nonlinear regression analysis and thereby calculate the MPN (Δ_{final}), and agree with conventional estimations, in a one parameter fit.

MPN from Nonlinear Regression

Numerical methods used to fit nonlinear functions to various empirical observations have been available for nearly 40 years (Hartley 1961). The most commonly used techniques for curve fitting, or nonlinear least squares approximations, are "linearization", "scoring", "steepest descent" and "Marquardt's compromise" (Afifi and Azen 1979; Draper and Smith 1980). We have chosen the modified Gauss-Newton linearization (Hartley 1961) approach because it is one of the most common techniques used for curve fitting in statistical software packages (Afifi and Azen 1979). To fully realize such a curve fitting-based MPN algorithm, we first manifest the method in a general form in order to derive our much simpler version. Thus, we collect Y_i (i = 1, 2, ..., I; for this work $Y_i = p_i$) data as a function of the controlled variable X_i (Φ_i) which we desire to fit to a model F_i (e.g., Eq. 5) dependent upon parameters Q_j (j = 1, 2, ..., J; for this work Q = Δ). For Q_j near Q_j^0, an initial guess, F_i is expanded about Q_j^0 in a "Taylor series". Using only first order terms,

$$F_i = F_i^0 + \sum_{j=1}^{J} \partial_{Q_j} F_i \, (Q_j - Q_j^0) = F_i^0 + \sum_{j=1}^{J} Z_{ij} \, \Delta Q_j^0 \tag{10}$$

and

$$F_i^0 = F_i - \sum_{j} Z_{ij} \, \Delta Q_j^0 \; ; \tag{11}$$

Z_{ij} are elements of the matrix **Z** (e.g., I rows × J columns as found in a spreadsheet)

$$\mathbf{Z} = \begin{bmatrix} \partial_{Q_1} F_1 & Z_{12} & \cdots & Z_{1J} \\ Z_{21} & \partial_{Q_2} F_2 & \cdots & Z_{2J} \\ \vdots & \vdots & \partial_{Q_j} F_i & \vdots \\ Z_{I1} & Z_{I2} & \cdots & \partial_{Q_J} F_I \end{bmatrix}. \tag{12}$$

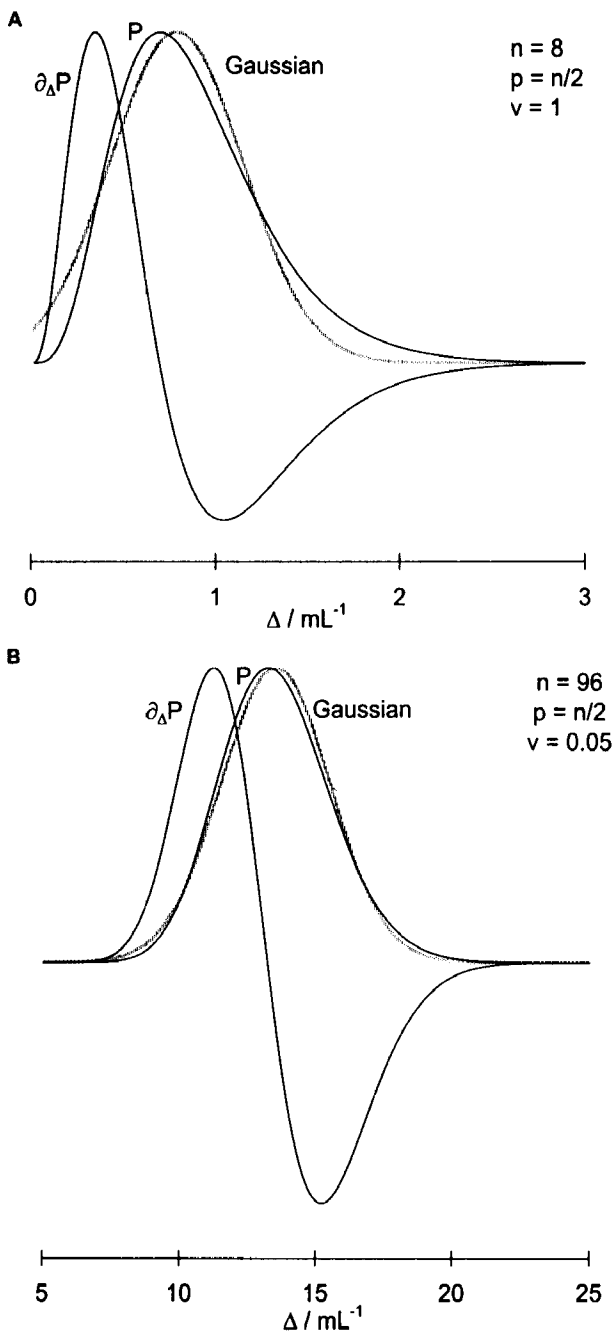


FIG. 3. PLOTS OF P, $\partial_\Delta P$, AND THE BEST FIT GAUSSIANS

The difference between Y_i and F_i calculated with Q_j^0 is

$$\Delta Y_i = Y_i - F_i^0 . \tag{13}$$

Combining Eq. 11 and 13, the error sum of squares is

$$ESS = \sum_{i=1}^{I} (\Delta Y_i)^2 . \tag{14}$$

To minimize, the partial first derivatives of ESS with respect to each j^{th} parameter (J derivatives) are set to zero, rearranged, and simplified such that for any particular parameter Q_φ

$$\sum_{i=1}^{I} Z_{\varphi i} \left(Y_i - F_i^0\right) = \sum_{j=1}^{J} \left[\sum_{i=1}^{I} \left(Z_{\varphi i} Z_{ij}\right) \Delta Q_j^0 \right] \tag{15a}$$

or, for all Q_j,

$$\mathbf{Z}' \begin{bmatrix} Y_1 - F_1^0 \\ Y_2 - F_2^0 \\ \vdots \\ Y_I - F_I^0 \end{bmatrix} = \mathbf{Z}'\mathbf{Z} \begin{bmatrix} Q_1 - Q_1^0 \\ Q_2 - Q_2^0 \\ \vdots \\ Q_J - Q_J^0 \end{bmatrix} = \mathbf{Z}'\mathbf{Z}\ \Delta\mathbf{Q}^0 . \tag{15b}$$

We seek a solution for $\Delta\mathbf{Q}^0$, a J×1 array, therefore each side of Eq. 15b is multiplied by the matrix inverse of $\mathbf{Z}'\mathbf{Z}$ ($\mathbf{Z}'$ is the transposition of $\mathbf{Z}$) to achieve

$$\Delta Q^0 = (\mathbf{Z}'\mathbf{Z})^{-1}\, \mathbf{Z}'(\mathbf{Y} - \mathbf{F}^0) \tag{16a}$$

Utilizing Excel format (evaluated using " Ctrl + Shift + Enter") and keeping in mind that $\mathbf{Z}$ {I × J}, $\mathbf{Y}$-$\mathbf{F}$ {I × 1}, and $\Delta\mathbf{Q}^0$ {J × 1} represent arrays of formulas,

$\Delta\mathbf{Q}^0$ = MMULT(MMULT(MINVERSE(MMULT(TRANSPOSE(**Z**),**Z**)), TRANSPOSE(**Z**)),**Y**–**F**). (16b)

In practice, a distinct $\Delta\mathbf{Q}^k$ array is generated at each k^{th} iteration and the appropriate j^{th} elements of $\Delta\mathbf{Q}^k$ *are arranged* to be automatically added to the various parameters used in the computation of $\Delta\mathbf{Q}^{k+1}$, etc. (Irwin *et al.* 1994, 2000). Eventually these results converge when $\Delta\mathbf{Q}^k$ no longer changes during successive calculations; ideally, upon convergence (k = K),

$$\Delta Q^{K} = \begin{bmatrix} \sim 0 \\ \sim 0 \\ \vdots \\ \sim 0 \end{bmatrix}. \tag{16c}$$

Applying Eq. 16a to Eq. 5 (e.g., $F_i = p_i = n - e^{-v\Delta\Phi i} n$, $\partial_p F_i = e^{-v\Delta\Phi i} nv\Phi i$, $Q = \Delta$, and dropping the superscript notation) simplifies to the straightforward evaluation

$$\Delta Q = \frac{\sum_i (Y_i - F_i)\,\partial_\Delta F_i}{\sum_i (\partial_\Delta F_i)^2}. \tag{17a}$$

In Fig. 4 (I = 12, J = 1) we solve for a new estimation of Δ (cell F4) as

$$\Delta_{new} = \Delta + \kappa \Delta Q \ . \tag{17b}$$

Using Excel's formulaic protocol, we input into cells F4, F6, {H2:H13}, and {D2:D13}, respectively

$$= F2 + 0.1 * F6 \,, \tag{17c}$$

$$=SUM((H2:H13)*(D2:D13))/SUM((D2:D13)^2) \,, \tag{17d}$$

$$=\{C2:C13\} - (\$E\$2*(1-EXP(-\$E\$4*\$F\$2*\{B2:B13\}))) \,, \tag{17e}$$

and

$$=(\{B2:B13\}*\$E\$2*\$E\$4)*EXP(-\{B2:B13\}*\$F\$2*\$E\$4) \,. \tag{17f}$$

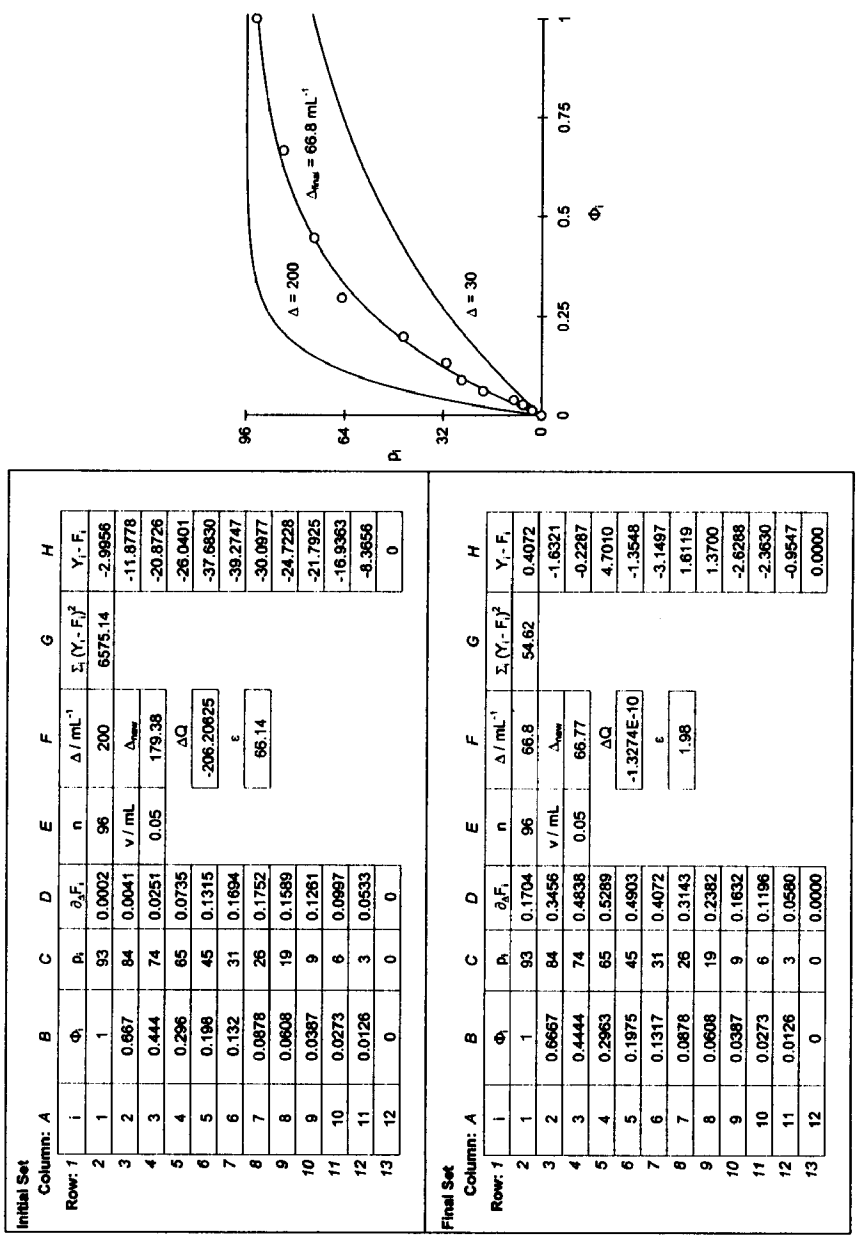
In Eq. 17d-f the symbols "{ }" bracket arrays of relative, or positional, references; for example, within cell H5, Eq. 17e appears as

"=C5-(E2*(1-EXP(-E4*F2*B5)))"

while in cell H11 it appears as

"=C11-(E2*(1-EXP(-E4*F2*B11)))".

Cell references given as "$LETTER$NUMBER" remain fixed regardless of position within an array. Equation 17d (cell F6) operates on data from several cells simultaneously and is only accurately evaluated when key-strokes "Ctrl + Shift + Enter" are used to "Enter" the expression. In Eq. 17c, cell F6 is multiplied by 0.1 (κ) in order to slow down the rate of change in ΔQ and thereby avoid problems overshooting the sought-after value of Δ (e.g., Δ is the MPN only with a minimized sum of squares; Eq. 13). Diminishing ΔQ in this way allows the use of very poor initial estimates for the MPN yet solve the problem quickly. For instance, when 2000 mL^{-1} was used in Fig. 4 as an initial estimate of Δ, the spreadsheet converged promptly (~ 160 iterations; ~ 1 s) even though ΔQ started at -5824. Lastly, by



Initial Set

Column:	A	B	C	D	E	F	G	H
Row: 1	i	Φ_i	p_i	$\partial_2 F_i$	n	Δ / mL^{-1}	$\Sigma_i (Y_i - F_i)^2$	$Y_i - F_i$
2	1	1	93	0.0002	96	200	6575.14	-2.9856
3	2	0.667	84	0.0041	v / mL	Δ_{new}		-11.8778
4	3	0.444	74	0.0251	0.05	179.38		-20.8726
5	4	0.296	65	0.0735		ΔQ		-26.0401
6	5	0.198	45	0.1315		-206.20625		-37.6830
7	6	0.132	31	0.1694		ε		-39.2747
8	7	0.0878	26	0.1752		66.14		-30.0977
9	8	0.0608	19	0.1589				-24.7228
10	9	0.0387	9	0.1261				-21.7925
11	10	0.0273	6	0.0997				-16.9363
12	11	0.0126	3	0.0533				-8.3656
13	12	0	0	0				0

Final Set

Column:	A	B	C	D	E	F	G	H
Row: 1	i	Φ_i	p_i	$\partial_2 F_i$	n	Δ / mL^{-1}	$\Sigma_i (Y_i - F_i)^2$	$Y_i - F_i$
2	1	1	93	0.1704	96	66.8	54.62	0.4072
3	2	0.6667	84	0.3456	v / mL	Δ_{new}		-1.6321
4	3	0.4444	74	0.4838	0.05	66.77		-0.2287
5	4	0.2963	65	0.5289		ΔQ		4.7010
6	5	0.1975	45	0.4903		-1.3274E-10		-1.3548
7	6	0.1317	31	0.4072		ε		-3.1497
8	7	0.0878	26	0.3143		1.98		1.6119
9	8	0.0608	19	0.2382				1.3700
10	9	0.0387	9	0.1632				-2.6288
11	10	0.0273	6	0.1196				-2.3630
12	11	0.0126	3	0.0580				-0.9547
13	12	0	0	0.0000				0.0000

FIG. 4. EXCEL SPREADSHEET ILLUSTRATING THE COMPUTATION FOR PERFORMING THE CURVE FITTING METHOD OF DETERMINING MPN

applying the theory of linear regression to the approximate linear equation we obtain an expression for the standard error for Δ_{final} (the "asymptotic standard error", Draper and Smith 1980; Irwin *et al.* 1994)

$$\varepsilon_{\Delta_{Final}} = \sqrt{\frac{\sum_i \left(Y_i - F_{i,Final}\right)^2}{\left(\sum_i \left(\partial_\Delta F_{i,Final}\right)^2\right) I - 1}} \quad . \tag{18a}$$

To calculate Eq. 18a (Fig. 4) with Excel we make cell F8 contain the formula

```
=SQRT(SUM((H2:H13)^2)/(SUM((D2:D13)^2)*
                    (COUNT(C2:C13)-1)))          (18b)
```

and evaluate the expression upon using keystrokes "Ctrl + Shift + Enter".

In Fig. 4 an initial guess (200 mL^{-1}) for Δ was introduced in cell F2 and the spreadsheet recalculated. Cell F2 was then made equal to F4 thereby creating the circular reference needed to start the iterative solving process and the calculation converged (~ 35 iterations) when the arrays ceased to change at the chosen numerical precision. Figure 4 also displays a plot of the binomial data p_i as a function of dilutions Φ_i as well as initial (shown for both 30 and 200 mL^{-1}) and final (~ 67 mL^{-1}) values of Δ. Regardless of the starting value for Δ, the calculation always converged to 66.8 ± 2.0 mL^{-1} which is close to (65.9 ± 3.6 mL^{-1}) the MPN from the maximum probability resolution protocol discussed in the Methods section.

Comparison of The MPN Protocols

Comparisons of the MPN methods (MPR and CF) using the same data sets are shown in Fig. 5 and Table 2. MPN values (20 experiments) from curve fitting ($\Delta_{final,CF}$, Eq. 17; $\kappa = 0.1$; $\pm \epsilon_\Delta$) are displayed plotted against those ($\Delta_{final,MPR}$, Eq. 9; $\pm \sigma_{MPR}$) based on the traditional MPR technique and show an almost perfect correlation (slope and r^2 of 0.99; Fig. 5). The inset scatter plot (Fig. 5) displays all the observed positive responses (p, out of n = 96) as a function of normalized dilution terms and demonstrates the high precision gained by performing the dilution method using large n. Quantitatively (Table 2), the two methods differed in MPN (or Δ_{final}) only by about 3% and displayed similar chi-squared (χ^2) statistics albeit the CF method was slightly better on average. While the agreement with total aerobic plate count (δ) was not perfect, the two MPN protocols averaged (absolute value) a modest 7% deviation with respect to δ; the CF approach was slightly better than the MPR method (Table 2). In general, the two MPN methods exhibited excellent agreement.

The curve fitting protocol which we have applied herein *should only be used* for large n (minimized rounding and sampling errors). To illustrate this concept we

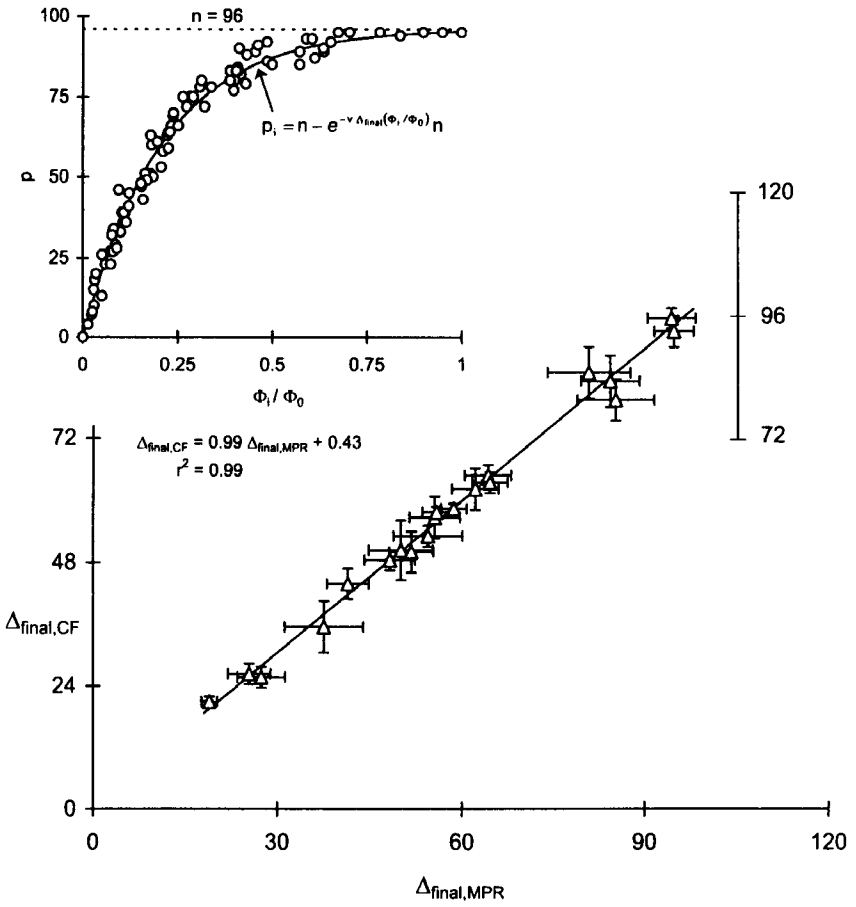


FIG. 5. COMPARISON OF CURVE FITTING (CF) AND MAXIMUM PROBABILITY RESOLUTION (MPR) DETERMINATIONS OF MPN ± THEIR RESPECTIVE ERROR TERMS

present comparisons of MPN values for the 3-dilution calculation with varying n (n = 5, n = 15, and n = 96; Table 3). For n = 15 and 96 the p_i-triads were made equivalent with respect to the MPN calculated from the n = 5 values therefore unlikely combinations were not presented; of course, improbable combinations of p_i are less likely to occur with greater n. The disagreement between the traditional MPN calculation (MPR, Table 3) and our curve fitting method (CF) averaged 238% for n = 5, ~ 2.46% for n = 15, and 0.288% (Table 2 deviations [absolute value] averaged ~ 3%) for n = 96. So, it would appear that the curve fitting MPN method could be utilized for n >15. However, for our micro-technique turbidity assay, n = 96 produced excellent results and a reasonably low limit of detection (Appendix).

TABLE 2.
COMPARISON OF *SALMONELLA ENTERITIDIS* ENUMERATION FROM PLATE COUNTING (δ) AND MPR/CF MPN PROTOCOLS (Δ) ACROSS 20 EXPERIMENTS

expt	δ [a]	$\frac{\Delta_{MPR}-\delta}{\Delta_{MPR}}$	$\frac{\Delta_{CF}-\delta}{\Delta_{CF}}$	MPR Δ [b]	σ_{MPR}	χ^2 [c]	CF Δ [d]	$\sigma_{CF} = \varepsilon$	χ^2	$\frac{\Delta_{MPR}-\Delta_{CF}}{\Delta_{MPR}}$	dilutions
1	21.0	-11%	0%	18.9	1	8	20.9	1	5	-11%	7
2	61.5	4%	5%	64.3	4	1	64.8	2	1	-1%	5
3	73.8	9%	13%	80.9	5	6	84.7	6	6	-5%	5
4	63.4	2%	0%	64.6	4	6	63.4	4	6	2%	5
5	84.6	0%	-2%	84.4	6	1	83.1	2	1	2%	5
6	94.2	0%	1%	94.4	6	3	95.3	4	3	-1%	5
7	89.4	6%	4%	94.9	6	3	92.7	5	3	2%	5
8	59.7	-2%	-2%	58.7	4	4	58.4	4	4	1%	5
9	57.6	-3%	0%	55.9	4	6	57.7	4	6	-3%	5
10	42.8	11%	12%	48.3	3	10	48.5	3	3	0%	4
11	43.4	16%	14%	51.8	4	2	50.4	2	2	3%	4
12	55.5	0%	2%	55.6	4	2	56.7	2	2	-2%	4
13	53.9	1%	-1%	54.5	4	4	53.1	2	3	3%	4
14	78.9	8%	1%	85.3	7	4	79.5	5	4	7%	4
15	37.4	1%	-6%	37.6	3	3	35.4	2	2	6%	4
16	66.3	-7%	-7%	62.2	5	6	62.2	5	6	0%	4
17	45.7	12%	9%	51.8	4	2	50.1	2	2	3%	4
18	38.7	7%	12%	41.5	3	5	43.8	3	4	-6%	4
19	30.2	-19%	-15%	25.4	2	1	26.3	1	1	-4%	4
20	33.6	-23%	-31%	27.4	2	2	25.7	1	1	6%	4
average of | x |:		7.06%	6.79%		4.05	3.91		2.99	3.26	3.30%	

[a] total aerobic plate count; average of 3 RA and 3 BHIA plates; CFU mL^{-1}

[b] maximum probability resolution (traditional) MPN method; mL^{-1}

[c] $\chi^2 = \sum_{i=1}^{k} \frac{\left(\left[n - e^{\lambda \bullet \Delta}\right]\left[n - p_i\right]\right)^2}{n\left(e^{\lambda \bullet \Delta_i^2} - 1\right)}$; Best 1990

[d] curve fitting MPN method; mL^{-1}

TABLE 3.
COMPARISON OF CALCULATED 3-DILUTION MPN ENUMERATION FROM THE MAXIMUM PROBABILITY RESOLUTION (MPR) AND CURVE FITTING (CF) METHODS AS A FUNCTION OF OBSERVATIONS PER DILUTION (n)

			MPN = Δ_{final} n = 5			Δ_{final} n = 15							Δ_{final} n = 96					
p_1[a]	p_2	p_3	*MPR*	*CF*	deviation	p_1	p_2	p_3	*MPR*	*CF*	deviation	p_1	p_2	p_3	*MPR*	*CF*	deviation
5	*3*	*4*	20.88	160.94	-670.79%	*15*	*13*	*3*	20.44	20.77	-1.61%	*96*	*84*	*18*	20.69	20.79	-0.48%
5	*3*	*3*	17.26	91.36	-429.32%	*15*	*12*	*2*	15.58	15.83	-1.60%	*96*	*79*	*15*	17.18	17.25	-0.41%
5	*3*	*2*	13.80	11.45	17.03%	*15*	*11*	*2*	13.25	13.32	-0.53%	*96*	*72*	*13*	13.92	13.94	-0.14%
4	*1*	*3*	3.05	1.92	37.05%	*14*	*4*	*0*	2.70	2.92	-8.15%	*92*	*26*	*3*	3.16	3.16	0.00%
4	*0*	*3*	2.29	1.48	35.37%	*14*	*3*	*0*	2.45	2.46	-0.41%	*88*	*21*	*2*	2.46	2.47	-0.41%
			ave. |deviation| :		238%					2.46%						0.29%

[a] $v = 1$, $\Phi_1 = 1$, $\Phi_2 = 0.1$, $\Phi_3 = 0.01$

Advantages of the Curve Fitting MPN Method

The CF method has several advantages over the MPR method described herein. It is much simpler to set up (Fig. 4) inasmuch as the direct method (I = 7) requires 1000 cells for each dilution i and an additional 1000 for the product (P_{total}). Secondly, the CF MPN calculation is faster. Less than 1 s (500 MHz, Pentium III; ca. 150 iterations s^{-1}) are required to calculate Δ_{final} even when the initial guess is off by several orders of magnitude. For instance, Fig. 6 shows values for Δ (calculated from the p_i-Φ_i pairs of Fig. 4) from the CF (Eq. 17; κ = 0.1; ~ 70 iterations) algorithm as a function of the number of iterative cycles. Even though the initial value used for the calculation was poor (1000 mL^{-1}) the CF method converged to 66-67 mL^{-1} in nearly 25 iterations. Contrariwise, the direct MPR method takes several minutes to set up (e.g., mainly determining the Δ_{start} and Δ_{end}, Fig. 2). A third advantage for performing our curve fitting MPN protocol over the traditional MPN method is that blank (diluent alone; p ~ 0) observations are used thereby gaining an additional degree of freedom. Lastly, since the curve fitting protocol is a correlation it has a more visual basis for *rejection* of data since excessive scatter is indicative of a dilution error. The closest standard for the rejection of traditional MPN data is the far less intuitive χ^2 statistic (Best 1990).

APPENDIX

One of the advantages of binomial methods is that they have very low limits of detection (LOD) relative to plate counting methods. Ordinarily, the limit of detection of any procedure is related to the intercept of a standard curve + an error term. However, binomial assays with a large n have an essentially zero intercept therefore the linear part (at the limit where the nonlinear function is approximately linear) of the relationship is

$$p = m\Delta \tag{A1}$$

$$m = \left.\frac{\partial p}{\partial \Delta}\right|_{\Delta \to 0} = n\,v \tag{A2}$$

therefore

$$\frac{1}{nv}\int_0^{p_{min}} dp = \int_0^{\Delta_{min}} d\Delta \tag{A3}$$

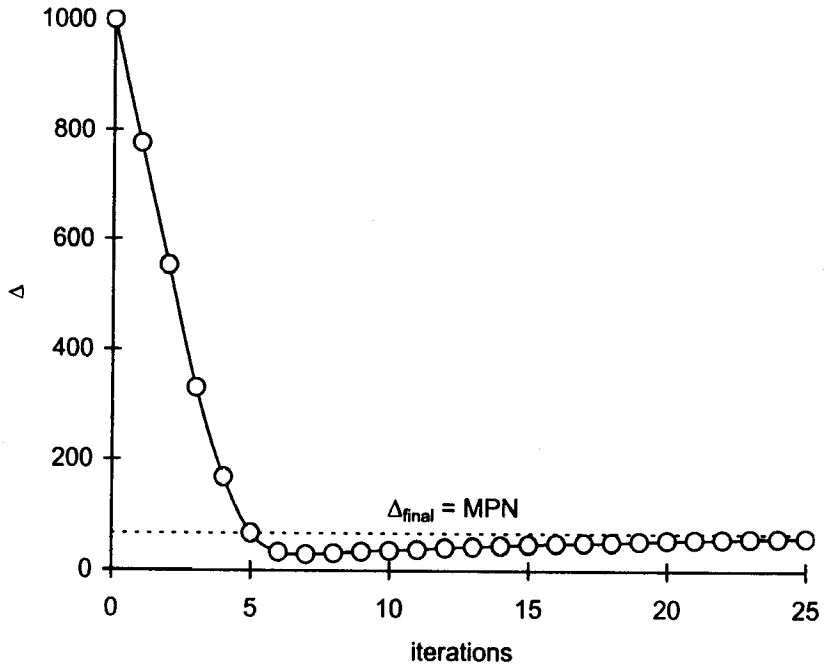


FIG. 6. ITERATIVE BEHAVIOR OF THE CURVE FITTING MPN METHOD WHEN STARTING FROM A VERY POOR GUESS AT Δ

$$\Delta_{min} = \frac{p_{min}}{nv}; \tag{A3}$$

Δ_{min} defines the limit of detection (LOD). For $p_{min} = 1$, $n = 96$, and $v = 0.05$ mL the 96-well turbidity assay has a LOD of ~ 0.2 mL^{-1}. However, across the numerous experiments (described herein) and taking into account the blank standard error (diluent alone),

$$\Delta_{min} = \frac{p_{min}}{nv} = \frac{1 + \left(\bar{p}_{blank} + \sigma_{\bar{p}_{blank}} t_{0.05}\right)}{nv} = 0.42\ mL^{-1}. \tag{A4}$$

Of course, the LOD could be improved by doubling v (and using more concentrated BHI). An identical LOD is obtained using a modified version of Eq. 8

$$\Delta_{min} = \frac{\mathrm{Log}_e \dfrac{n}{n - \left(1 + \left(\bar{p}_{blank} + \sigma_{\bar{p}_{blank}} t_{0.05}\right)\right)}}{v}. \tag{A5}$$

REFERENCES

AFIFI, A.A. and AZEN, S.P. 1979. Regression and correlation analysis. In *Statistical Analyses: A Computer Oriented Approach*, 2nd Ed., Academic Press, New York.

BARKWORTH, H. and IRWIN, J.O. 1938. Distribution of coliform organisms in milk and the accuracy of the presumptive coliform test. J. Hygiene *38*, 446-457.

BEST, D.J. 1990. Optimal determination of most probable numbers. Intl. J. Food Microbiol. *11*, 159-166.

BLAIS, B.W. and YAMAZAKI, H. 1991. Application of polymyxin-coated polyester cloth to the semi-quantitation of *Salmonella* in processed foods. Intl. J. Food Microbiol. *14*, 43-50.

BRIONES JR., A.M. and REICHARDT, W. 1999. Estimating microbial population counts by 'most probable number' using Microsoft Excel. J. Microbiol. Methods *35*, 157-161.

COCHRAN, W.G. 1950. Estimation of bacterial densities by means of the "most probable number". Biometrics June, 105-116.

DAVIS, H.F. 1973. Differential forms, matrices, and tensors. In *Introduction to Vector Analysis*, 2nd Ed., Allyn and Bacon, Boston.

DE MAN, J.C. 1975. The probability of most probable numbers. Europ. J. Applied Microbiol. *1*, 67-78.

DE MAN, J.C. 1983. MPN tables, corrected. Europ. J. Applied Microbiol. Biotechnol. *17*, 301-305.

DICKSON, J.S. 1989. Enumeration of *Salmonellae* by most-probable-number using the *Salmonella* 1-2 test. J. Food Protec. *52*, 388-391.

DRAPER, N.R. and SMITH, H. 1980. An introduction to nonlinear estimation. In *Applied Regression Analysis*, 2nd Ed., Wiley, New York.

FUNG, D.Y.C. and KRAFT, A.A. 1969. Rapid evaluation of viable cell counts by using the microtiter system and MPN technique. J. Milk Food Technol. *32*, 408-409.

HAINES, J.R., WRENN, B.A., HOLDER, E.L., STROHMEIER, K.L., HERRINGTON, R.T. and VENOSA, A.D. 1996. Measurement of hydrocarbon-degrading microbial populations by a 96-well plate most-probable-number procedure. J. Industrial Microbiol. *16*, 36-41.

HALVORSON, H.O. and ZIEGLER, N.R. 1933. Application of statistics to problems in bacteriology. I. A means of determining bacterial population by the dilution method. J. Bacteriology *25*, 101-121.

HARTLEY, H.O. 1961. The modified Gauss-Newton method for the fitting of nonlinear regression functions by least squares. Technometrics *3*, 269-280.

HUMBERT, F.S., SALVAT, G., LALANDE, F. and COLIN, P. 1997. Miniaturized most probable number and enrichment serology technique for the enumeration of *Salmonella* spp. on poultry carcasses. J. Food Protect. *60*, 1306-1311.

IRWIN, P.L., BROUILLETTE, J.N., GERMANN, M., HICKS, K.B., KURANTZ, M. and DAMERT, W. 1994. Calculation of immobilized enzyme reaction progress curves from nested ordered-sequential rate expressions. Enzyme Microbial Technol. *24*, 675-686.

IRWIN, P.L., DAMERT, W. and DONER, L. 1994. Curve fitting in nuclear magnetic resonance spectroscopy: illustrative examples using a spreadsheet and microcomputer. Concepts Magnetic Resonance *6*, 57-67.

IRWIN P.L., GEHRING, A., TU, S., BREWSTER, J., FANELLI, J. and EHRENFELD, E. 2000. Minimum detectable level of *salmonellae* using a binomial-based bacterial ice nucleation detection assay. J. AOAC Intl. *83*, 1087-1095.

McCRADY, M.H. 1915. The numerical interpretation of fermentation-tube results. J. Infect. Dis. *17*, 183-212.

ROWE, R., TODD, R. and WAIDE, J. 1977. Microtechnique for most-probable-number analysis. Applied Environ. Microbiol. *33*, 675-680.

STEEL, R.G.D. and TORRIE, J.H. 1960. Enumeration data *I*: one-way classification. In *Principles and Procedures of Statistics,* McGraw-Hill Book Co., New York.

TURPIN, P.E., MAYCROFT, K.A., BEDFORD, J., ROWLANDS, C.L. and WELLINGTON, E.M.H. 1993. A rapid luminescent-phage based MPN method for the enumeration of *Salmonella typhimurium* in environmental samples. Letts. Applied Microbiol. *16*, 24-27.